Journal of Alexandria University for Administrative Sciences© Vol. 59, No. 2, March 2022

Mok-Nskart: New Procedure for estimating Confidence interval of Steady State Simulation Output Mean[1]

Dr. Somaia Mohamed Said

Lecturer of applied statistics, Department of Statistics, Mathematics& Insurance

Faculty of Commerce, Alexandria University, Egypt.

somaia.said@alexu.edu.eg

ABSTRACT

Discrete-event simulation can be classified as either finite-horizon or steady-state. Finite-horizon (terminating) simulation models ended at a specific time or by the occurrence of a specific condition. On the other hand, steady-state (non-terminating) simulation operates (at least conceptually) into the indefinite future; and in this case the interest focus on long-run average performance.

Non-overlapping Batch Means (NBM) method is one of the most effective methods used for analyzing steady-state simulation output.

This paper presents a simulation procedure undertaken in the automotive industry. Therefor it aims First, to develop a new procedure for steady-state simulation output analysis "Mok-Nskart", it is designed to deliver a confidence interval for the steady-state mean of a simulation output process. Mok-Nskart can be considered as an extension of the method of NBM. The present study aims secondly, to evaluate the performance of Mok-Nskart, by comparing it with other simulation analysis procedures—namely N-Skart and MSER-5.

Keywords: *Confidence interval (CI), Non-overlapping Batch Means (NBM), Steady state simulation analysis.*

[1] Received in 15/12 /2021, accepted in 25/1/2022.

1. INTRODUCTION

The complexity of many real-world systems involves many unaffordable analytical models: queuing systems, supply chain and sustainability. consequently, such systems are commonly studied using simulation. Simulation is useful to measure performance in systems that are so complex that they cannot be described by analytical queuing models (Abubakar; Adamu; Abdulkadir; Abdulkadir, 2020; Hanna; Abdelghany; Abdou, 2021; Elhariry, 2020 & Elgazzar, 2021). Common model taxonomy classifies a simulation problem along three main dimensions (Law & Kelton, 2000): (i) Deterministic vs. Stochastic. (ii). Static vs. Dynamic (depending on whether they require a time component). (iii). Continuous vs. Discrete (depending on how the system changes). Discrete event Simulation is a specific technique for modeling stochastic and or deterministic, dynamic, and evolving system.

Discrete Event Simulation is a comprehensive tool for the analysis and design of manufacturing systems. Over the years, considerable efforts to improve simulation processes have been made. One step in these efforts is the standardization of the output data through the development of an appropriate system which presents the results in a standardized way (Barrera; Oscarsson; Lidberg & Sellgren, 2018).

Simulation software is one of common manager's aids that facilitates the modeling and calculation of more complicated situations (Al-Fedhly and Folley, 2020).

Three fundamental problems arise in analyzing output from a stochastic steady-state simulation. The first problem is initialization bias, as it is usually impossible to start a simulation in steady-state operation. The second problem is autocorrelation; caused by pronounced stochastic dependencies among successive responses generated within a single simulation run. The third problem is the non-normality problem caused by pronounced departures from normality in the simulation-generated responses.

These problems prevent the construction of valid confidence intervals (CI) s for the steady state mean, that is because standard statistical methods require independent identically distributed (i.i.d.) normal observations to yield a valid CI.

A good CI procedure requires dealing with the previous three problems to provide the following:

- An accurate point estimator for the steady state mean that is approximately free of initialization bias.
- A sufficiently stable estimator of the standard error of the point estimator that adequately takes into consideration any correlations between the simulation responses used in computing the point estimator.
- A suitable adjustment to the usual critical value of Student's t -distribution that adequately accounts for any departures from normality in the simulation responses used in computing the point estimator and the standard error estimator.

The method of batch-means seeks to obtain a sequence of independent samples (batch-means) by aggregating *h* successive observations of a steady state simulation. The batch size must be large enough so that the first order correlation of batch-means is below a positive small threshold.

In the NBM method, the sequence of simulation-generated outputs $\{y_i : i = 1,2, \ldots, n\}$ is divided into *k* adjacent non overlapping batches, each of size *m*. For simplicity, it is assumed that *n* is a multiple of *m* so that *n* = *km*. The sample mean for the *j*th batch is:

$$\bar{y}_j(\mathrm{m}) = \frac{1}{m} \sum_{i=(j-1)m+1}^{jm} y_i \qquad \forall\, j = 1,2,\ldots,\mathrm{k} \qquad (1)$$

The grand mean $\bar{y}(n)$ of the individual batch means, is used as a point estimator for the steady state mean γ

$$\bar{y}(n) = \bar{\bar{y}}(m,k) = \frac{1}{k}\sum_{j=1}^{k} \bar{y}_j(m) = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (2)$$

The objective is to construct a CI estimator for γ that is centered on a point estimator as in Equation (2).

If the batch size *m* is sufficiently large so that the batch means $\{\bar{y}_j(\mathrm{m}): j = 1, 2, \ldots, k\}$ are approximately independent and identically distributed (i.i.d.) normal random variables with mean γ, then classical results concerning Student's *t*-distribution can be applied. The sample variance of the k batch means for batches of size m is $s^2(m,k)$

$$s^2(m,k) = \frac{1}{k-1}\sum_{j=1}^{k}\left[\bar{y}_j(m) - \bar{\bar{y}}(m,k)\right]^2 \qquad (3)$$

If the original simulation-generated process $\{y_i : i = 1, 2, \ldots, n\}$is stationary and weakly dependent, then it follows that as $m\to\infty$ with k fixed so that $n\to\infty$, an asymptotically valid $100(1-\alpha)$ % CI for γ is:

$$\left[\bar{\bar{y}}(m,k) \pm t_{\left(k-1,1-\frac{\alpha}{2}\right)}\sqrt{\frac{s^2(m,k)}{k}}\right] \qquad (4)$$

Conventional NBM procedures such as ABATCH and LBATCH (Fishman and Yarberry, 1997) and the procedure of Law and Carson (Law and Carson, 1979) are based on Equation (4); and they are designed to determine the batch size m and the batch count k that are required to satisfy approximately the assumption of i.i.d. normal batch means. If this assumption is satisfied exactly, then we obtain a CI whose actual coverage probability is exactly equal to the nominal level $1-\alpha$.

By contrast, the more recent NBM procedures ASAP Steiger; Lada; Wilson; Alexopoulos; Goldsman and Zouaoui (2002); ASAP3 Steiger; Lada; Wilson; Joines; Alexopoulos and Goldsman (2005); SBatch Lada; Steiger and Wilson (2008) and Skart Tafazzoli (2009), are designed to determine sufficiently large values of the batch size and the length of the initial warm-up period so as to ensure that batch means computed beyond the warm-up period are approximately multivariate normal with identically distributed marginal (that is, they approximately constitute a stationary Gaussian process) but are not necessarily independent.

If the resulting batch means are correlated, then the classical NBM *t*-ratio underlying Equation (4) does not possess Student's *t*-distribution with $k-1$ degrees of freedom, so that an appropriate modification of Equation (4) is required to yield an approximately valid CI for γ.

In this article Mok-Nskart is presented, it is a non-sequential procedure for steady-state simulation analysis which is an extension of the classical method of batch means. Mok-Nskart is intended for simulation experiments in which the size of the output data set is fixed because of a limited computing budget, a constraint on the time available for the user to complete the simulation study, or other restrictions that prevent the user from resuming the current run of the simulation model. Mok-Nskart is designed to deliver a CI for the steady-state mean that has a user-specified coverage probability and that is based on a single time series of an arbitrary fixed length. Mok-Nskart can be considered as an extension of both two recent procedures N-Skart (Tafazzoli; Lada; Steiger; Wilson; 2011) and MSER-5 (Mokashi, 2010).

The rest of this article is organized as follows. Section 2 provides a brief overview of N-Skart and MSER-5 procedures. Section 3 contains an overview of the new procedure (Mok-Nskart).

Section4 presents selected results from the experimental performance evaluation. Section 5 contains the conclusions and recommendations for future work.

2. OVERVIEW OF PROCEDURES TO BE COMPARED

2.1 Overview of MSER-5

Starting a model running from an 'unrealistic' state can lead to the occurrence of initialization bias. This causes the output data collected at and near the beginning of the run to be uncharacteristic of the later and 'true' output steady state value. If this uncharacteristic data is included in the calculation of the overall response value, it can produce a biased result and therefore incorrect conclusions. One method for dealing with this problem is to run the model for a warm-up period until steady state is reached and remove the initialization bias by deleting the data within that warm-up period (Hoad &Robinson, 2011).

MSER-5 was proposed by Franklin & White (2008); Mokashi (2010) and white; Cobb and Spratt (2000), to deliver an improved simulation-based point estimator of the steady-state mean. It aims to balance improved accuracy achieved by reducing the estimator's bias against the loss of precision (increased variance) caused by truncating the original data set through deletion of some leading observations. It is based on the following rationale given a time series $\{y_i : i = 1, \ldots, n\}$ of simulation-generated responses having fixed length (sample size) N which will be used to compute an accurate estimator of the steady-state mean

γ, and to find a data-truncation point beyond which all the remaining observations are typical of steady-state behavior. For each candidate data-truncation point in the data set, MSER-5 computes a CI for γ based on all the observations beyond the truncation point; and takes the half-length of this CI as a measure of the extent to which all the remaining observations are typical of steady-state behavior, where a smaller CI half-length indicates closer conformity to steady-state behavior. It follows that the data-truncation truncation point (that is, the length of the warm-up period) should be set to minimize the length of the CI for γ based on the remaining (truncated) output sequence.

MSER-5 operates on non-overlapping batch means (NBM) with batch size 5 to ensure more stable behavior in the CIs used to determine the truncation point; MSER-5 constrains the data-truncation point to be in the first half of the given data set. Therefore, MSER-5 uses as its basic data items the batch means with batch size 5

$$y_j = \frac{1}{5}\sum_{i=1}^{5} y_{5(j-1)+i} \ , \quad j = 1,2,\dots,k = \left\lfloor \frac{N}{5} \right\rfloor \qquad (5)$$

Where, for each real number u, the floor function $\lfloor u \rfloor$ denotes the greatest integer not exceeding u.

If d denotes the data-truncation point (that is, the length of the warm-up period), then the grand average $\bar{y}(k,d)$ and sample variance $s_y^2(k,d)$ of the truncated batch means $y_j \ : \ j = d+1,\dots,k$ are:

$$\bar{y}(k,d) = \frac{1}{k-d}\sum_{j=d+1}^{k} y_j \quad , d = 0,1,2,\dots,\left\lfloor \frac{k}{2} \right\rfloor \qquad (6)$$

$$s_y^2(k,d) = \frac{1}{k-d}\sum_{j=d+1}^{k} \left(y_j - \bar{y}(k,d)\right)^2 \qquad (7)$$

Respectively and 100(1 – α) % CI for γ has the form

$$\bar{y}(k,d) \pm z_{1-\frac{\alpha}{2}} \frac{s_y(k,d)}{\sqrt{k-d}} \qquad (8)$$

White et al. (2000) do not recommend using (8) as the final CI estimator for γ; instead, they recommend merely using (8) as a device for determining the optimal truncation point d^* as follows:

$$d^* = arg\ min_{0 \le d < \lfloor \frac{k}{2} \rfloor} \frac{z_{1-\frac{\alpha}{2}} s_y(k,d)}{\sqrt{k-d}} \qquad (9)$$

But if $d^* = \lfloor \frac{k}{2} \rfloor$, then MSER-5 fails because of inadequate sample size.

Unfortunately, when $d^* = \lfloor \frac{k}{2} \rfloor$ so that MSER-5 fails to deliver any estimator of γ, white et al. (2000) do not suggest a method for increasing the sample size so that MSER-5 can ultimately deliver the desired point and CI estimators of γ .

If $d^* < \lfloor \frac{k}{2} \rfloor$, then MSER-5 delivers the truncated sample mean $\bar{y}(k, d^*)$ as the final point estimator of the steady-state mean. To compute the associated nominal 100(1 – α) % CI for γ , MSER-5 applies the classical NBM method to the truncated sequence y_j : $j = d^* + 1,2,\dots,k$, which is now regarded as the "original" (raw, unbatched) observations from which It is possible to compute:

$k^* = 20$ "new" batch means with batch size $m^* = \lfloor \frac{k-d^*}{k^*} \rfloor$

Therefore, the ι th new batch mean is computed as:

$$\bar{y}_\iota(m^*, d^*) = \frac{1}{m^*}\sum_{i=1}^{m^*} y_{d^*+m^*(\iota-1)+i}\ , \quad \iota = 1,2,\dots,k^* \qquad (10)$$

And the corresponding grand average and sample variance of the new batch means are given respectively by:

$$\bar{\bar{y}}_\iota(k^*, m^*, d^*) = \frac{1}{k^*}\sum_{\iota=1}^{k^*} \bar{y}_\iota(m^*, d^*) \qquad (11)$$

$$s_{\bar{y}}^2(k^*, m^*, d^*) = \frac{1}{k^*-1}\sum_{\iota=1}^{k^*} [\bar{y}_\iota(m^*, d^*) - \bar{\bar{y}}_\iota(k^*, m^*, d^*)]^2 \qquad (12)$$

The final 100(1 − α) % CI for γ is:

$$\bar{y}(k, d^*) \pm t_{\left(k^*-1,1-\frac{\alpha}{2}\right)} \frac{s_{\bar{y}}(k^*, m^*, d^*)}{\sqrt{k^*}} \tag{13}$$

Where $t_{\left(k^*-1,1-\frac{\alpha}{2}\right)}$ denotes the $1-\frac{\alpha}{2}$ quantile of Student's t-distribution with k^*-1 degrees of freedom.

CI in (13) is consistent with the recommendations of (Schmeiser, 1982) on applying the method of non-overlapping batch means to a data set $y_j : \; j = d^* + 1,2, \dots, k$ that which is a realization of a covariance stationary simulation output process.

2.2 Overview of N-Skart

N-Skart was proposed by Tafazzoli et al. (2011), the input to N-Skart is a simulation-generated time series $\{y_i : i = 1, \dots, n\}$ of fixed length N, where $N \geq 1280$ and the user specifies the required coverage probability $(1-\alpha)$ (where $0 < \alpha < 1$) for a CI estimator of γ based on the given data set. N-Skart handles the start-up problem by applying the randomness test of (Von Neumann, 1941) to determine sufficiently large values of the batch size m and spacer size dm (where m ≥ 1 and d ≥ 0) such that the corresponding k spaced batch means:

$$y_j(m, d) = \frac{1}{m} \sum_{i=1}^{m} y_{\{j(d+1)-1\}m+i} \;, j = 1,2, \dots, k \tag{14}$$

are approximately independent of each other and of the initial condition. It follows that any effects due to initialization bias are limited to the initial spacer $\{y_i : i = 1, \dots, dm\}$; and this is the reason why N-Skart uses the initial spacer as the warm-up period so that the first dm observations are deleted (ignored).

If the first dm observations -that are deleted- are so large, then the data set size N may be not large enough to enable N-Skart to determine sufficiently large values for the spacer size and batch size such that the spaced batch means pass the randomness test, then N-Skart issues a warning and gives the user options either to stop or to continue anyway in computing point and CI estimators of γ .

Beyond the truncation point dm, N-Skart computes $\grave{k}$ truncated, nonspaced batch means with batch size m

$$y_j(m) = \frac{1}{m}\sum_{i=1}^{m} y_{(d+j-1)m+i} \quad , j = 1,2,\ldots,\grave{k}$$

Where $\grave{k}$ is taken large enough to use as much of the data set $\{y_i : i = 1, \ldots, n\}$ as possible; and then N-Skart computes the sample mean $\bar{y}(m,\grave{k})$ and variance of the truncated non-spaced batch means $s^2_{m,\grave{k}}$ respectively,

$$\bar{y}(m,\grave{k}) = \frac{1}{\grave{k}}\sum_{j=1}^{\grave{k}} y_j(m) \quad , s^2_{m,\grave{k}} = \frac{1}{\grave{k}}\sum_{j=1}^{\grave{k}}\left[y_j(m) - \bar{y}(m,\grave{k})\right]^2 \quad (15)$$

Finally, N-Skart delivers an asymptotically valid 100(1 − α) % skewness- and autocorrelation-adjusted CI for γ having the form:

$$\bar{y}(m,\grave{k}) - G(L)\sqrt{\frac{A\, s^2_{m,\grave{k}}}{\grave{k}}}, \bar{y}(m,\grave{k}) - G(R)\sqrt{\frac{A\, s^2_{m,\grave{k}}}{\grave{k}}}) \qquad (16)$$

Where the skewness adjustments G(L) and G(R) are defined in terms of the function G(ω) ,where:

$$G(\omega) \equiv \frac{\sqrt[3]{1+6\beta(\omega-\beta)}-1}{2\beta} \quad ; \ \beta = \frac{\hat{B}}{6\sqrt{\ddot{k}}} \qquad (17)$$

$\hat{B}$ is approximately unbiased estimator of the marginal skewness of $y_j(m)$ computed from the $\ddot{k}$ spaced batch means of the form (14) with batch size m that are separated by spacers of size at least dm. $\ddot{k}$ is taken large enough to use the entire data set of size N.

The skewness-adjustment function G (.) has the arguments:

$R = t_{\frac{\alpha}{2},\ddot{k}-1}$; $\qquad L = t_{1-\frac{\alpha}{2},\ddot{k}-1}$

and the correlation adjustment A is computed as:

$$A = \frac{1 + \hat{\varphi}_{y(m)}}{1 - \hat{\varphi}_{y(m)}} \qquad (18)$$

Where $\hat{\varphi}_{y(m)}$ is the standard estimator of the lag-one correlation of the truncated non-spaced batch means.

$$\hat{\varphi}_{y(m)} = corr\left[y_j(m), y_{j+1}(m)\right]$$

So, N-Skart is dealing with the problem of initialization bias by some of inaccuracies. Whereas to determine warm-up period N-Skart applies the randomness test of (Von Neumann, 1941) to the current set of batch means to determine

the required batch count, batch size, and data-truncation point beyond which all computed spaced batch means are approximately independent and identically distributed (i.i.d.). While that may require increasing batch size or the spacer size in each time the test is failed, and that may be required to increase the whole sample size over than the size of simulation run N.

In this case, N-Skart issues a warning to the user, stating that the randomness test could not be passed because of insufficient data. The warning also notes that if the user decides to continue the procedure under the given circumstances, then the delivered CI might not achieve the target confidence level. Here the user has two choices: a) quit the procedure without delivering a CI; or b) continue with construction of the requested CI by ignoring the warning. Therefore, it does not determine warm-up period with sufficient accuracy, resulting in inaccurate estimate of CI for γ.

This problem has been avoided in MSER-5 as it deals with the problem of initialization bias differently. However, it deals with this problem only and does not deal with the problems of autocorrelation and data skewness.

Therefore, pursuant to the recommendation by Mokashi; Tejada; Yousefi; Xu; Wilson and Tafazzoli (2010), the current paper suggests a new procedure based on a combination of both N-Skart and MSER-5 procedures, taking the advantages of both in estimating CI for γ. Researchers have launched the proposed Procedure name "Mok-Nskart".

3. OVERVIEW OF THE NEW PROCEDURE (MOK-NSKART)

Mok-Nskart relies first on MSER-5 in identifying data truncation point and dealing with the problem of initialization bias, a way that ensures accurate estimate CI for γ. Then Mok-Nskart relies on N-Skart in dealing with the problems of autocorrelation and skewness, and finally constructs a valid estimate CI for γ . Fig. 1 depicts a high-level flowchart of Mok-Nskart.

3.1 Detailed Algorithmic Statement of Mok-Nskart

The first two steps are like MSER-5 procedure, while the rest of the steps are like N-skart.

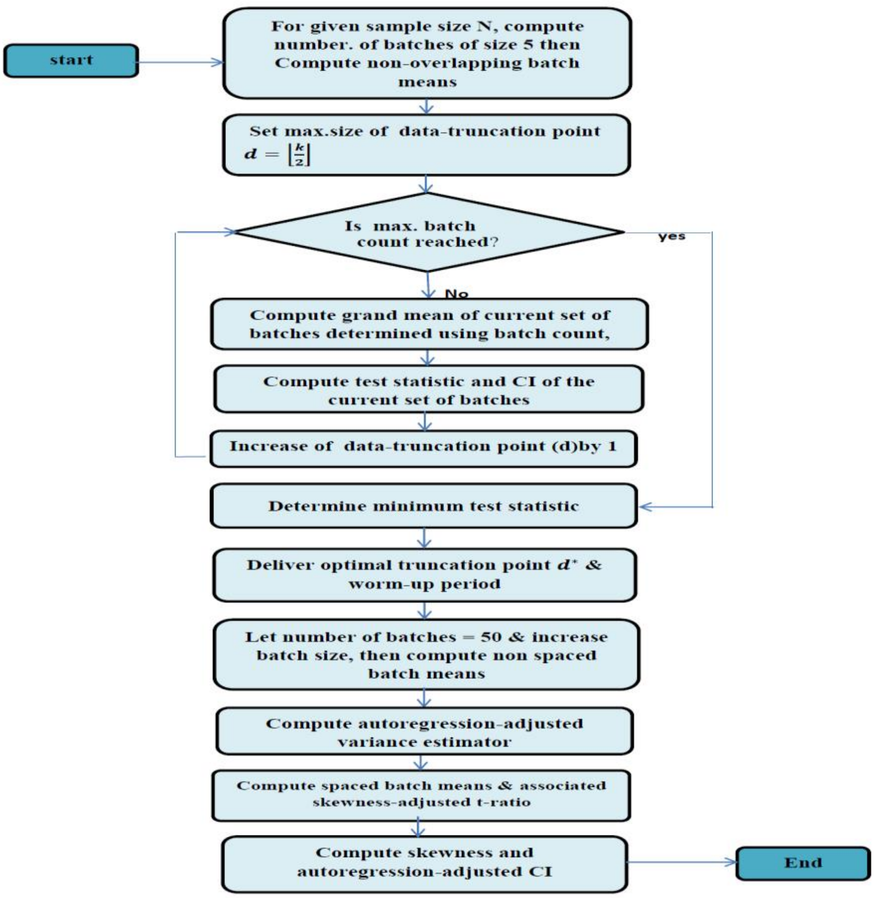

Fig 1: High-level flowchart of Mok-Nskart

– **Data preparation**

Divide the given sample data set of size N to k batches, each of them has size m=5, such that $k = \left\lfloor \frac{N}{m} \right\rfloor$ .Compute batch means y_j in Equation (5) and use the batch means with batch size 5 as basic data items. Then compute the grand average $\bar{y}(k,d)$ in Equation (6) and variance of truncated batch means $s_{\bar{y}}^2(k,d)$ in Equation (7)

– **Warm-up period**

Skip the first d batches from the batch means data set where $\{d = 0,1,2,\dots,\left\lfloor\frac{k}{2}\right\rfloor\}$, then use the truncated batch means to compute a CI analog to each value of d as in (8).

Depending on (8), compute the CI half-length analog to each value of d, the optimal truncation point d^* will be that value at which CI half-length is minimum. Therefore, warm-up period has ($d^* * m$) observations. Where,

$$d^* = arg\,min_{0\le d<\left\lfloor\frac{k}{2}\right\rfloor} \frac{z_{1-\frac{\alpha}{2}} S_y(k,d)}{\sqrt{k-d}} \qquad (9)$$

– **Autocorrelation Adjustment for the Variance Estimator**

Use the truncated means - after skipping d^* - as the basic data. Divide it into $k^* = 50$ batches with size $m^* = \left\lfloor\frac{k-d^*}{k^*}\right\rfloor$. In this way, number of batches in Mok-N-Skart exceeds those in MSER-5 (starting with $k^* = 21$,then increased k^* up till results became stable), batch size m^* in Mok-Nskart also is bigger than that in N-skart. In that manner Mok-Nskart forms a set of approximately independent batch means.

Compute both of: truncated batch means $\bar{y}_l(m^*,d^*)$in Equation (10), grand mean $\bar{\bar{y}}_l(k^*,m^*,d^*)$ in Equation (11) and sample variance $s_{\bar{y}}^2(k^*,m^*,d^*)$ in Equation (12).

Depending on equation (12), $\frac{s_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}$ can be used as an unbiased estimator of variance of the grand mean $var[\bar{\bar{y}}_l(k^*,m^*,d^*)]$ under assuming that means are i.i.d. However, means are rarely uncorrelated, therefore $\frac{s_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}$ is considered as a biased estimator to variance of the grand mean.

So, Mok-Nskart is like N-Skart adjusts $\frac{s_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}$ by the correlation-adjustment factor A based on (18) considering the effect of autocorrelation in batch means. Therefore, $\frac{A\, s_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}$ will be an unbiased estimator of $var[\bar{\bar{y}}_l(m^*,d^*)]$. It is worth noting, Mok-Nskart depends on equation (11) as a point estimator computed from truncated means, contrary to what has been followed in MSER-5 which considered the truncated sample mean $\bar{y}(k, d^*)$ in (6) as a point estimator.

So, Mok-Nskart is like N-skart, adjusts $\frac{s^2_{\bar{y}}(k^*,m^*,d^*)}{k^*}$ by the correlation-adjustment factor A to compensate for any residual correlation between the truncated batch means; that will improve the performance of Mok-Nskart's final CIs (Tafazzoli et al., 2011).

– **Skewness Adjustment to Student's t-Statistic**

When the truncated, non-spaced batch means exhibit significant departures from normality, Mok-Nskart -similar to N-skart- applies an appropriate adjustment to the usual critical value of Student's t -distribution to yield a valid CI for γ (Tafazzoli et al., 2011). Mok-Nskart inflates the batch size for a highly skewed process to mitigate at least partially the effect of non-normality of the batch means on the associated NBM Student's t –statistic.

Therefore, the skewness adjustment that Mok-Nskart applies in this step can be crucial in delivering a CI with good coverage. Moreover, it has been found that the batch-size increases imposed in previous steps of Mok-Nskart are necessary to ensure that the skewness of the batch means has sufficiently small magnitude, so the skewness adjustment is effective. In this step, Mok-Nskart adapted the skewness adjustment developed by (Willink, 2005 & 2006).

To obtain an approximately unbiased estimator of the marginal skewness of the current set of truncated, nonspaced batch means, Mok-Nskart computes this skewness estimator from approximately i.i.d. spaced batch means, constituting a subset of the current set of non-spaced batch means. The associated spaced batch means have approximately unbiased estimators of their required marginal moments: the grand mean in equation (11), the sample variance in equation (12) and the sample third central moment ζ(k^*,m^*,d^*), where:

$$\zeta(k^*, m^*, d^*) = \frac{k^*}{(k^*-1)(k^*-2)} \sum_{\iota=1}^{k^*} \left(\bar{y}_\iota(m^*, d^*) - \bar{\bar{y}}(k^*, m^*, d^*)\right)^3 \quad (19)$$

An approximately unbiased estimator of the marginal skewness of truncated batch means is $\hat{B}$, where:

$$\widehat{\mathrm{B}} = \frac{\zeta(k^*, m^*, d^*)}{s_{\bar{y}}^3(k^*, m^*, d^*)} \tag{20}$$

– **The correlation and Skewness-adjusted CI**

Calculate $\mathrm{G(L)}$ and $\mathrm{G(R)}$, the skewness-adjusted quantile of Student's t - ratio for the left and right half-lengths of the proposed CI. The function $\mathrm{G}(\omega)$is defined by (17). Thus Mok-Nskart provides the correlation and skewness-adjusted CI

$$\left(\bar{\bar{y}}(k^*, m^*, d^*) - G(L)\sqrt{\frac{As_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}}, \bar{\bar{y}}(k^*, m^*, d^*) - G(R)\sqrt{\frac{As_{\bar{y}}^2(k^*,m^*,d^*)}{k^*}}\right)$$

Finally, Mok-Nskart delivers skewness- and autoregression-adjusted $100(1-\alpha)$ % CI for γ with the user-specified coverage probability and stops.

4. PERFORMANCE EVALUATION OF MOK-NSKART

To evaluate the performance of Mok-Nskart with respect to coverage probability and the mean and variance of the half-length of its CIs, Mok-Nskart was applied to a carefully selected set of test problems as explained in (Tafazzoli, 2009)& (Tafazzoli et al., 2011) including processes with characteristics that are typical of many large-scale practical applications of steady-state simulation, and processes exhibiting extremes of stochastic behavior that are commonly used to stress-test simulation analysis procedures. For each test process, the steady-state mean is known; therefore, the empirical coverage probabilities were employed for the CIs delivered by Mok-Nskart as the primary means of evaluating the performance of the procedure. To illustrate the performance of Mok-Nskart for what might be considered "small," "medium," and "large" data sets, the following sample sizes in experiments are used: 10 000; 20 000; 100 000; and 200 000.

Some results are also presented for N-Skart and MSER-5, providing a direct comparison with the performance of Mok-Nskart. Beyond CI coverage probability, the performance of Mok-Nskart is reported with respect to the following criteria: average CI relative precision (that is, the CI's half-length expressed as a percentage of the magnitude of the CI's midpoint); average CI half-length; and variance of the CI half-length. Each experiment includes 1000 independent replications of the selected output analysis procedures for constructing 90% and 95% CIs. Given below is a brief description of each of the test processes used in our performance evaluation of Mok-Nskart. Complete details for each of these test processes are given in (Tafazzoli, 2009) & (Tafazzoli et al., 2011).

- **(M/M/1) Queue Waiting Time Process with Empty Initial Condition and 90% Server Utilization**

The test process is the sequence of waiting times in the queue with an empty-and-idle initial condition, an interarrival rate of $\lambda = 0.9$ customers per time unit, and a service rate of $\mu = 1$ customers per time unit. In this system the steady-state server utilization is $\rho = 0.9$, and the steady-state expected waiting time is $\gamma = \frac{\rho}{\mu(1-\rho)} = 9$ time units.

–**(M/M/1) Queue Waiting Time Process with Empty Initial Condition and 80% Server Utilization:**

The test process is defined in the same way as for the previous test process, except the interarrival rate is $\lambda = 0.8$ customers per time unit so that the steady-state server utilization is $\rho = 0.8$, and the steady-state expected waiting time is $\gamma = 4$ time units.

- **(M/M/1) Queue Waiting Time Process with 113 Initial Customers and 90% Utilization**

This test process has the same interarrival rate and service rate as the first test process (λ=0.9 and μ=1), but with an additional condition that 113 customers are already present in queue at time 0. The first "regular" customer arrives after time 0 and begins service after the initial 113 customers have finished service. This initial condition ensures that the expected queue waiting time for the first regular customer to arrive after time 0 is 10 steady-state standard deviations above the steady-state mean, ensuring a pronounced initial transient for this test process as explained in (Tafazzoli, 2009).This process has the same steady-state parameters as the first test process—i.e., a server utilization of ρ=0.9 and steady-state mean queue-waiting time of γ=9 time units

- **M/H2/1 Queue Waiting Times**

The test process is the sequence of waiting times in a queuing system with an empty-and-idle initial condition, a mean interarrival time of 1.0, and a hyperexponential service-time distribution that is a mixture of two exponential distributions such that the service times have a mean of 0.8 and a coefficient of variation of 2.0. Thus, in steady-state operation this system has a server utilization of $\rho = 0.8$ and a mean queue-waiting-time of $\gamma = 8$ time units.

– **First-Order Autoregressive (AR (1)) Process**

The test process is an AR (1) process with autoregressive parameter 0.995 and white-noise variance 1.0, so that the steady-state distribution of the process has marginal mean $\gamma = 100$ and marginal standard deviation $\sigma_X = 10.0125$. Because the initial condition $\tilde{y}_0 = 0$ was chosen, this test process has pronounced negative initialization bias. Although this process is normal, it has a pronounced autocorrelation structure that severely distorts the behavior of the classical method of batch means.

The M/M/1 queue-waiting-time process is characterized by a relatively short warm-up period. However, the process exhibits a pronounced autocorrelation structure, with the autocorrelation function for the waiting time decaying slowly as the lag increases. Also, the M/M/1 queue waiting times have a steady-state probability distribution which has a non-zero probability mass at zero and an exponential tail. This results in a slow convergence of the batch means to the normal distribution with increasing batch size.

Depending on R language, version i386 2.15.2 and RStudio 0.98.1091 platform a programming code is written to perform simulation and get results for our selected models.

Table 1 summarizes the performance of Mok-Nskart, MSER-5and N-Skart on the selected M/M/1 queue-waiting-time process.

From the results in Table 1,

– It is evident that the CI properties obtained from N-Skart were better than those obtained from Mok-Nskart and MSER-5. The CI overages delivered by N-Skart were close to the corresponding nominal coverage levels. CIs delivered by Mok-Nskart had acceptable overages, in the sense that Mok-Nskart gives better results than MSER-5, and less performance than of N-skart. However, it is virtually overcoming the shortcomings of both procedures.

– For smaller sample sizes, it can be said that the performance of N-Skart to be unacceptable, N-Skart failed in all iterations (Failed = 1000 times) to pass Von Neumann test, which is used to determine warm-up period, and then CIs delivered by it will be influenced by initialization bias. Moreover, these CIs will be based on sample size less than that required to pass randomness test, resulting in wider half lengths and misleading coverages.

Table 1

Performance Of Mok-Nskart, N-Skart And MSER-5 For M/M/1 Queue-waiting Time Process with $\rho = 0.9$ And empty – and Idle initial condition Based On 1000 Replications of Each Procedure

Overall sample size N	Procedure	90% Confidence Interval Properties							95%Confidence Interval Properties						
		coverage	Rel.prec.	Avg.half	Var.half	Bias	MSE	Failed	coverage	Rel.prec.	Avg.half	Var.half	Bias	MSE	Failed
10000	N-skart	88%	31.25%	3.18	5.25	0.08	3.46	1000	91.7%	37.86%	4.02	7.24	0.01	3.84	1000
	Mser-5	67%	22.85%	1.97	0.79	0.57	3.23	0	69.3%	28.08%	2.46	1.51	0.48	4.02	0
	Mok-Nskart	76.4%	27.43%	2.59	2.95	0.64	3.33	0	80.1%	33.47%	3.36	7.69	0.59	4.18	0
20000	N-skart	88.8%	23.49%	2.26	1.08	0.06	1.71	946	94.8%	28.5%	2.83	1.87	0.00	1.81	961
	Mser-5	73.3%	18.9%	1.67	0.51	0.33	1.96	0	79.4%	22.94%	2.03	0.64	0.32	1.84	0
	Mok-Nskart	79.6%	21.09%	2	1.41	0.39	2.00	0	87.3%	25.96%	2.6	3.06	0.37	1.88	0
50000	N-skart	90.8%	16.81%	1.63	0.71	0.01	0.75	311	96.4%	20.23%	2	1.49	0.06	0.69	287
	Mser-5	80.2%	13.84%	1.24	0.20	0.12	0.80	0	87.1%	16.63%	1.49	0.31	0.17	0.71	0
	Mok-Nskart	82.3%	14.88%	1.43	0.88	0.14	0.80	0	90.6%	18.33%	1.84	2.54	0.19	0.72	0
100000	N-skart	92.7%	11.97%	1.13	0.28	0.04	0.34	6	95.8%	14.42%	1.37	0.46	0.07	0.38	7
	Mser-5	84.7%	10.36%	0.93	0.07	0.08	0.34	0	89.4%	12.64%	1.13	0.11	0.11	0.39	0
	Mok-Nskart	85.6%	10.78%	1	0.25	0.09	0.34	0	89.6%	13.47%	1.29	0.61	0.11	0.39	0
200000	N-skart	92%	8.3%	0.77	0.03	0.03	0.17	0	95.9%	10.02%	0.93	0.07	0.01	0.18	0
	Mser-5	87.6%	7.62%	0.68	0.03	0.06	0.17	0	91.7%	9.14%	0.82	0.04	0.04	0.18	0
	Mok-Nskart	89%	7.81%	0.72	0.07	0.06	0.17	0	92.1%	9.39%	0.87	0.13	0.04	0.18	0

– As the sample size increased, the CI coverages delivered by Mok-Nskart were close to the nominal coverage level, as they were also close to the coverage delivered by N-skart; for example with the sample size N =200,000, the nominal 90% CIs delivered by N-Skart had an empirical coverage probability of 92% and an average relative precision of 8.3%, while the corresponding measures by Mok-Nskart were 89% and 7.81% respectively, whereas they were 87.6% and 7.62% by MSER-5.

– The CIs delivered by MSER-5 are the less dispersion at all, the variance of the CI half-lengths is the smallest –compared with Mok-Nskart & N-Skart – at all sample sizes and different confidence levels.

- With respect to the point estimators of γ delivered by Mok-Nskart, MSER-5 and N-Skart, it was observed that for all the sample sizes considered, the point-estimator bias was substantially larger for both Mok-Nskart and MSER-5 than for N-Skart.

Figure 2 displays the distribution of CIs half-lengths delivered by all three procedures in this test process for the selected sample sizes. It is clear from both histogram and boxplot analog to each procedure that CIs delivered by N-Skart and Mok-Nskart had half-lengths skewed to the right, with mode, which is greater than that of MSER-5. Meaning that CIs delivered by N-Skart and Mok-Nskart were the wider, therefore, their actual coverage was the best.

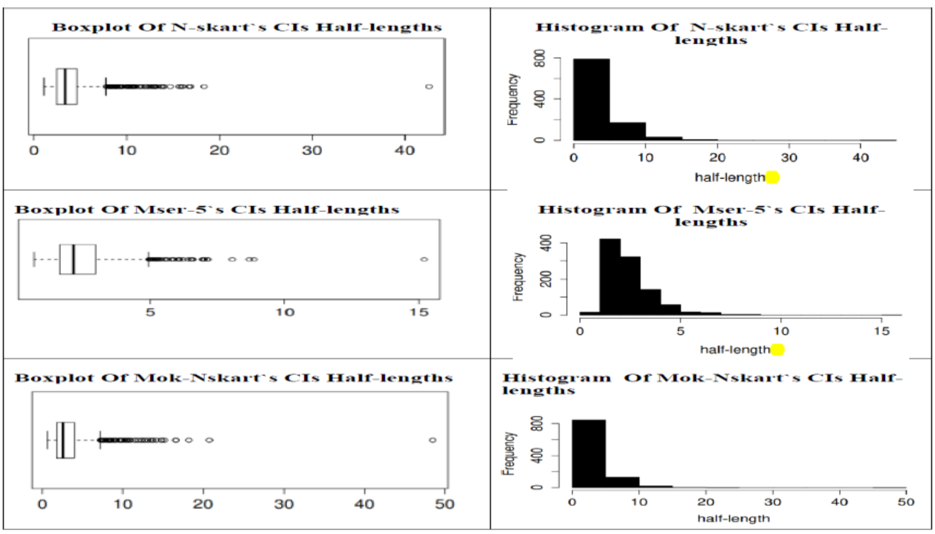

Figure 2

Boxplot and Histogram of CIs Half-lengths delivered by N-skart, MSER-5 And Mok-Nskart For M/M/1 Queue-waiting Time Process With $\rho = 0.9, N = 10000$ *and* $\alpha = 0.05$ And Empty – And Idle Initial Condition

5. CONCLUSION

In this article a new, completely automated batch-means procedure was developed, called Mok-Nskart. It considered as a combination of both N-Skart and MSER-5 procedures, taking the advantages of both in estimating CI for steady state mean γ . It works for constructing a correlation-and skewness-adjusted CI for the steady-state mean of a simulation output process in which the user supplies a single simulation-generated series of arbitrary length, and the user specifies the desired coverage probability for a CI based on that series. from the experimental results presented in Section 4.

It is evident that Mok-Nskart provided close conformance to the user-specified CI coverage probabilities in all the test problems that was considered.

The study recommends depending on Mok-Nskart procedure in simulating some other processes, such as M/M/1/LIFO, M/M/1/SIRO and (s, S) inventory control systems. Also, it recommends depending on Mok-Nskart to estimate a confidence interval for the steady state variance.

References

Abubakar, A. M.; Adamu, A.; Abdulkadir, A. and Abdulkadir H. S. (2020). Discrete Event Simulation of Clients Flow in Ante-natal Clinic. *Asian Journal of Probability and Statistics*, 6(2), 63-78.

Al-Fedhly, H. and Folley, D. (2020). Discrete Event Simulation Software for Agent-Based Supply Chain Demand. *International Journal of Industry and Sustainable Development,* 1(2), 20-35.

Barrera, C. A.; Oscarsson, J.; Lidberg, S., & Sellgren, T. (2018). Discrete Event Simulation Output Data-Handling System in an automotive Manufacturing plant. *8th Swedish Production Symposium, SPS, 16-18, Stockholm, Sweden.*

Elgazzar, S.H. (2021). investigating Digital Supply Chain Transformation Drivers: An Empirical Study. *Journal of Alexandria University of administrative sciences,* 58(5), 139-170.

Elhariry, B. M. (2020). A Structural Model for Measuring the Impact of Entrepreneurial Orientation on Sustainable Banking Performance: An Empirical Study on the Egyptian Banking Sector. *Journal of Alexandria University of administrative sciences,* 57(3), 55-98.

Fishman, G. S. and Yarberry, L. S. (1997). An Implementation of The Batch Means Method. *INFORMS Journal on Computing,* 9 (3), 296–310.

Franklin, W.W. and White, K. P. (2008). Stationarity Tests and MSER-5: Exploring the Intuition Behind Mean-Squared-Error Reduction in Detecting and Correcting Initialization Bias. In Proceedings of the 2008 Winter Simulation Conference, S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds 541-546.

Hanna, H.; Abdelghany, M. and Abdou, S. (2021). Developing a Sustainable Supply Chain Management Practices Framework: A Case Study on the Egyptian Wheat Sector. *Journal of Alexandria University of administrative sciences,* 58(2), 243-259.

Hoad, K., & Robinson, S. (2011). Implementing MSER-5 in Commercial Simulation Software and its Wider Implications. In Proceedings of the 2011 Winter Simulation Conference, S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, eds. 495-503.

Lada, E. K.; Steiger, N. M. and Wilson, J. R. (2008). SBatch: A Spaced Batch Means procedure for Steady-State Simulation Analysis. Journal of Simulation.2, 170-185 [accessed June 20, 2008] *http://www.ise.ncsu.edu /jwilson/files/lada08jos.pdf*

Law, A. M. and Carson, J.S. (1979). A sequential Procedure for Determining the Length of A steady State Simulation. *Operations Research,* 27(5), 1011–1025.

Law, A. M. and Kelton, W. D. (2000). Simulation modeling and analysis. Third Edition, McGraw-Hill, New York.

Mokashi, A. C. (2010). The Simulation Start-up Problem: Performance Comparison of N-Skart and MSER-5. Master's thesis, North Carolina State University, Raleigh, North Carolina. [accessed July 16, 2010] *www.ise.ncsu.edu/jwilson/files/mokashi10ms. pdf*

Mokashi, A. C.; Tejada, J. J.; Yousefi, S.; Xu, T.; Wilson, J.R. and Tafazzoli, A. (2010). Performance Comparison of N-Skart and MSER-5on the Simulation Start-up Problem. In Proceedings of the 2010 Winter Simulation Conference, B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, eds. 971-982.

Schmeiser, B. W. (1982). Batch size effects in the analysis of simulation output. *Operations Research,* (30), 556-568.

Steiger, N. M. and Wilson J.R. (1999). Improved Batching for Confidence Interval Construction in Steady-State Simulation. In Proceedings of the 1999 Winter Simulation Conference. P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds. 442–451.

http://www.informs-cs.org/wsc99papers/061.PDF.

Steiger, N. M.; Lada E. K.; Wilson, J. R.; Alexopoulos, C.; Goldsman, D. and Zouaoui, F. (2002). ASAP2: An Improved Batch Means Procedure for Simulation Output Analysis, In Proceedings of the 2002 Winter Simulation Conference. E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds. 336-344.

Steiger, N. M.; Lada, E.K.; Wilson, J.R.; Joines, J.,A. Alexopoulos, C. and Goldsman, D. (2005). ASAP3: A Batch Means Procedure for Steady-State Simulation Analysis. *ACM Transactions on Modeling and Computer Simulation* 15(1), 39–73.

Tafazzoli, A. (2009). Skart: A Skewness- and Autoregression-Adjusted Batch-Means Procedure for Simulation Analysis, *Ph.D. dissertation*, Dept. Indust. Syst. Eng., North Carolina State Univ., Raleigh, NC.

www.lib.ncsu.edu/theses/available/etd-01122009-53054/unrestricted

Tafazzoli, A.; Lada, E.K.; Steiger, N. M. and Wilson, J. R. (2011). N-Skart: A Non-Sequential Skewness and Autoregression-Adjusted Batch-Means Procedure for Simulation Analysis, *IEEE Trans.* 56(2), 254-264

Von Neumann, J. (1941). Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *The Annals of Mathematical Statistics,* (12), 367–395

White, K. P. and Robinson, S. (2009). The Problem of the Initial Transient (again), or why MSER works. In Proceedings of the 2009 INFORMS. L.H. Lee, M.E. Kuhl, J.W. Fowler, and S. Robinson, eds. 90-95.

www.informs-sim.org/2009informs-simworkshop/paper92-97.pdf

White, K. P.; Cobb, M. J. and Spratt, S. C. (2000). A Comparison of Five Steady-State Truncation Heuristics for Simulation. In Proceedings of the 2000 Winter Simulation Conference, J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds. 755-760.

Willink, R. (2005). A Confidence Interval and Test for the Mean of an Asymmetric Distribution. *Communications in Statistics—Theory and Methods,* (34), 753–766.

Willink, R. (2007). On a Confidence Interval for the Mean of an Asymmetric Distribution. *Communications in Statistics—Theory and Methods*. 36 (12), 2235 -2236.

Mok-Nskart: أسلوب جديد لتقدير فترة الثقة للمتوسط من مخرجات المحاكاة في حالة الاستقرار

د. سمية محمد سعيد

ملخص البحث باللغة العربية

عادة ما ينصب الاهتمام فى المحاكاة غير المنتهية على خطوة التحليل الإحصائي للعملية العشوائية التى تمثل مخرجات المحاكاة، والسعى إلى الوصول لتقديرات نقطة وفترات ثقة للمتوسط فى وضع الاستقرار الإحصائى. إلا أنه، غالباً ما تكون هذه المخرجات بعيدة بشكل كبير عن وضع الاستقرار الإحصائى، كما أنها قد تكون مرتبطة ذاتياً ولا تتبع التوزيع الطبيعى. وهذا بدوره يؤدى إلى عدم إمكانية الاعتماد على إحصائية Student's t التقليدية عند تقدير فترات الثقة. من أجل ذلك، تقترح الدراسة الحالية أسلوب جديد لتقدير فترة ثقة للمتوسط فى وضع الاستقرار الاحصائى من مخرجات المحاكاة في حالة الاستقرار.

وقد سعت هذه الدراسة إلى تحقيق هدفين رئيسين، تمثل الأول في التغلب على أوجه القصور التى عانت منها بعض الأساليب الآلية السابقة، من خلال تصميم أسلوب آلى مقترح من قبل الباحثة لتحليل مخرجات المحاكاة في حالة الاستقرار (أسلوب Mok-nskart). أما الهدف الثاني فقد تمثل في قياس قدرة الأسلوب المقترح Mok-nskart على تقدير فترة ثقة للمتوسط في وضع الاستقرار γ، وعقد مقارنات بين أداء الأساليب الثلاث Mok-nskart و N-Skart و MSER-5 من خلال التطبيق على بعض أنظمة خطوط الانتظار وغيرها.

الكلمات الدالة: فترة الثقة CI، طريقة متوسطات الحزم غير المتداخلة NBM، المحاكاة في حالة الاستقرار.

Suggested Citation according to APA Style

Said, S. M. (2022). Mok-Nskart: New Procedure for estimating Confidence interval of Steady State Simulation Output Mean. *Journal of Alexandria University for Administrative Sciences*, *Faculty of Commerce*, *Alexandria University* 59(2), 115- 137.